



**University of
Zurich^{UZH}**

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2020

Microviridin 1777: A Toxic Chymotrypsin Inhibitor Discovered by a Metabologenomic Approach

Sieber, Simon ; Grendelmeier, Simone M ; Harris, Lonnie A ; Mitchell, Douglas A ; Gademann, Karl

Abstract: The toxicity of the cyanobacterium *Microcystis aeruginosa* EAWAG 127a was evaluated against the sensitive grazer *Thamnocephalus platyurus*, and the extract possessed strong activity. To investigate the compounds responsible for cytotoxicity, a series of peptides from this cyanobacterium were studied using a combined genomic and molecular networking approach. The results led to the isolation, structure elucidation, and biological evaluation of microviridin 1777, which represents the most potent chymotrypsin inhibitor characterized from this family of peptides to date. Furthermore, the biosynthetic gene clusters of microviridin, anabaenopeptin, aeruginosin, and piricyclamide were located in the producing organism, and six additional natural products were identified by tandem mass spectrometry analyses. These results highlight the potential of modern techniques for the identification of natural products, demonstrate the ecological role of protease inhibitors produced by cyanobacteria, and raise ramifications concerning the presence of novel, yet uncharacterized, toxin families in cyanobacteria beyond microcystin.

DOI: <https://doi.org/10.1021/acs.jnatprod.9b00986>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-196838>

Journal Article

Accepted Version

Originally published at:

Sieber, Simon; Grendelmeier, Simone M; Harris, Lonnie A; Mitchell, Douglas A; Gademann, Karl (2020). Microviridin 1777: A Toxic Chymotrypsin Inhibitor Discovered by a Metabologenomic Approach. *Journal of natural products*, 83(2):438-446.

DOI: <https://doi.org/10.1021/acs.jnatprod.9b00986>

Microviridin 1777: A Toxic Chymotrypsin Inhibitor Discovered by a Metabologenic Approach

Simon Sieber,^{†,‡} Simone M. Grendelmeier,^{†,‡} Lonnie A. Harris,[§] Douglas A. Mitchell,[§]
and Karl Gademann^{*,†}

[†]Department of Chemistry, University of Zurich, Winterthurerstrasse 190, Zurich CH 8057,
Switzerland

[§]Department of Chemistry, University of Illinois at Urbana–Champaign, 600 South Mathews
Avenue, Urbana, Illinois 61801, United States

[‡]These authors contributed equally

* correspondence to karl.gademann@uzh.ch

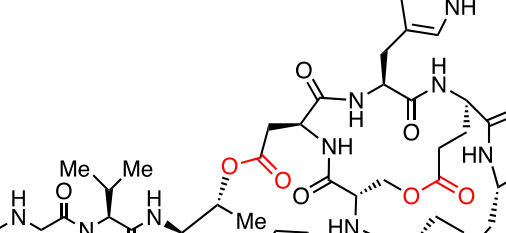
ABSTRACT

The toxicity of the cyanobacterium *Microcystis aeruginosa* EAWAG 127a was evaluated against the sensitive grazer *Thamnocephalus platyurus* and the extract possessed strong activity. To investigate the compounds responsible for cytotoxicity, a series of peptides from this cyanobacterium were studied using a combined genomic and molecular networking approach. The results led to the isolation, structure elucidation, and biological evaluation of microviridin 1777, which represents the most potent chymotrypsin inhibitor characterized from this family of peptides to date. Furthermore, the biosynthetic gene clusters of microviridin, anabaenopeptin, aeruginosin, and piricyclamide were located in the producing organism, and six additional natural products were identified by tandem mass spectrometry analyses. These results highlight the potential of modern techniques for the identification of natural products, demonstrate the ecological role of protease inhibitors produced by cyanobacteria, and raise ramifications concerning the presence of novel, yet uncharacterized, toxin families in cyanobacteria beyond microcystin.

Natural products remain an important source of biologically active compounds for drug discovery.^{1,2} The field has advanced tremendously over the recent years by exploiting innovative strategies such as genome-mining and molecular networking.³⁻⁵ These approaches allow for the identification of previously unknown structures, even in producing strains that have been thoroughly investigated by classical methods. Using genome-mining and networking approaches, gene clusters of natural products can be predicted by online tools (antiSMASH,⁶⁻⁹ PRISM,¹⁰ BAGEL¹¹⁻¹³ and RODEO¹⁴⁻¹⁶) and the compounds subsequently identified by the Global Natural Products Social Molecular Networking (GNPS) algorithm.⁴ The addition of other tools to the GNPS platform, such as VarQuest¹⁷ and DEREPLICATOR+,¹⁸ facilitate the detection of known natural products and the prediction of their analogues. The large and diversified database has been applied to the identification and discovery of natural products from various organisms,¹⁹⁻²³ exemplified by the isolation of the novel lipopeptides poaeamide B and bananamides 1, 2 and 3.²⁴ These research tools have been particularly successful in the analysis of cyanobacterial extracts²⁵⁻²⁸ recently highlighted by the global metabolomics study of *Microcystis*.²⁹

In the context of our work on new cyanobacterial toxins,³⁰⁻³⁶ we have postulated that strong protease inhibitors can display biological effects beyond deterrence leading to fatal effects on grazers. Experimental support for this hypothesis is documented in a number of studies,³⁷⁻⁴⁰ providing evidence that peptides from different classes such as cyanopeptolins and aeruginosins can display strong grazer toxicity.^{31-34,41} In order to further corroborate the production of multiple toxins by cyanobacteria, we were interested in potent protease inhibitors from other classes, which have not yet been characterized as cyanobacterial toxins. Building on these previous studies, we initiated a screening campaign for the identification of natural products possessing strong cytotoxicity using *Thamnocephalus platyurus* (*T. platyurus*) as a grazer model system. The results

A



B

ω -ester

ω -ester

Ac-Tyr¹-Asn²-Val³-Thr⁴-Leu⁵-Lys⁶-Tyr⁷-Pro⁸-Ser⁹-Asp¹⁰-Trp¹¹-Glu¹²-Glu¹³-Phe¹⁴

ω -amide

Results and Discussion

A preliminary screening targeted at grazer toxicity and cytotoxicity was performed evaluating several cyanobacterial extracts on *T. platyurus*. Extracts from *Microcystis aeruginosa* EAWAG 127a (*M. aeruginosa* EAWAG 127a) were found to exhibit modest cytotoxicity against *T. platyurus* (LD₅₀ = 0.43 mg/mL, Supporting Information, **Figure S12**). Building on the previously discussed hypothesis that potent protease inhibitors can be lethal to grazers, we decided to study the secondary metabolite peptides produced by *M. aeruginosa* EAWAG 127a. Additionally, the interest in this cyanobacterial strain was reinforced by its promising results from earlier screening.⁴² A combination of metagenomic and metabolite profiling was used for this study. A

small-scale extraction followed by UHPLC-HRMS/HRMS measurements were performed and the data were analyzed using the GNPS platform. Furthermore, the DNA of the organism was extracted following a standard protocol,⁴³ sequenced, and the resulting data were analyzed using the genome-mining tools RODEO¹⁴⁻¹⁶ and antiSMASH.⁶⁻⁹

To identify the peptides produced by *M. aeruginosa* EAWAG 127a, biosynthetic gene clusters of (ribosomally synthesized and post-translationally modified peptides) (RiPPs) and non-ribosomal peptides (NRP) were identified by antiSMASH and the produced natural products were identified using the GNPS platform. Using this approach, the microviridin-like biosynthetic gene cluster was identified. This cluster bears two precursor peptides, which were confidently predicted by virtue of their canonical TxKxPSD motif⁴⁴⁻⁴⁷ (**Figure 2** and Supporting Information, **Table S5**). The two compounds were identified as microviridin 1777 (**1**)⁴⁸⁻⁵⁰ and microviridin O.^{51,52} Microviridins have been discovered and recognized as protease inhibitors since 1990.⁵³ They are exemplified by a tricyclic scaffold and their biosynthesis has been extensively studied.^{46,47,50,54-56} These compounds constitute a class of RiPP⁵⁰ and the intramolecular ω -ester and ω -amide bonds are formed by ATP-grasp enzymes (members of InterPro families IPR026439 and IPR026446).^{46,47}

The analysis of the extract by UHPLC revealed that microviridin 1777 (**1**) was found as the major compound in the *M. aeruginosa* EAWAG 127a extract. As the microviridin class of compounds has been characterized for its serine protease inhibitory activity,⁴⁴ we hypothesized that the compound was primarily responsible for the observed *T. platyurus* toxicity. We therefore isolated, characterized, and evaluated the biological activity purified microviridin 1777 (**1**).

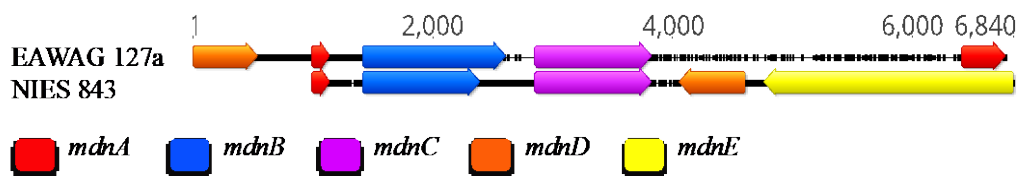


Figure 2. Alignment of the microviridin B and 1777 biosynthetic gene clusters. Accession ID (microviridin B): AM943877.1. From left to right: (protein accession IDs) CAQ_16116.1 to CAQ_16120.1.⁴⁴ Accession ID: (microviridin 1777): SRLN000000000. From left to right: (protein accession IDs): EZJ55_03500 to EZJ55_03525. *mdnA*, precursor peptide (red); *mdnB* and *mdnC*, ATP-grasp ligases (blue, purple); *mdnD*, *N*-acetyltransferase (orange); *mdnE*, ABC transporter (yellow).⁴⁴

In order to obtain a sufficient quantity of microviridin 1777 (**1**), a larger scale fermentation (50 L) of *M. aeruginosa* EAWAG 127a was initiated. The isolation process was optimized by pre-separation on a solid-phase extraction column using mixtures of MeOH and H₂O, which successfully separated another compound possessing similar retention time on the UHPLC column. Microviridin 1777 (**1**) was subsequently isolated by multiple preparative HPLC runs and its structure was elucidated (**Figure 3**) by extensive NMR (**Table 1**) and HRMS/HRMS analyses.

The molecular formula of the compound was calculated to be C₈₇H₁₁₁N₁₇O₂₄ by the detection of the doubly protonated molecule [M + 2H]²⁺ = 889.9058 Da. Based on ¹H, ¹³C NMR (**Table 1**), and the UV spectrum (**Supporting Information, Figure S2**), the compound was deduced to be a peptide, which supported the genomic data.²⁹ Each amino acid was identified by their typical proton and carbon chemical shifts and their correlation in COSY, TOCSY, HMBC and ROESY spectra.^{37,53,57-}

⁶⁰ The connectivity between the residues in the side chain was determined to be Ac-Tyr¹-Asn²-Val³ by the analysis of the key correlations on the HMBC, ROESY and TOCSY spectra. The acetyl group was deduced to be on the tyrosine (Tyr¹) residue due to the HMBC correlation between the NH of the amino acid and the carbonyl of the acetyl group. Following the same strategy, the

connection between Tyr¹ to Asn², Asn² to Val³, and Val³ to Thr⁴, was determined by the HMBC correlation between the NH atom of one amino acid with the carbonyl C atom of the adjacent residue. The amino acid sequence of the first cycle was elucidated to be Thr⁴-Leu⁵-Lys⁶-Tyr⁷-Pro⁸-Ser⁹-Asp¹⁰ using the HMBC correlation between the NH to the C=O of the other amino acid except for Pro⁸, Ser⁹ and Asp¹⁰. The connectivity between Pro⁸ and Tyr⁷ was deduced with a ROESY correlation between the aromatic proton of Tyr⁷ and the α -proton of Pro⁸, the connectivity between Pro⁸ and Ser⁹ was determined by the HMBC correlation between the α -proton of Ser⁹ to the C=O of Pro⁸. No correlations connecting Ser⁹ and Asp¹⁰ were observed. The composition of the N-terminal macrolactone was supported by the HMBC correlation between Thr⁴ H-3 and the γ -C=O of Asp¹⁰. The amino acids forming the second macrocycle were identified as Trp¹¹ and Glu¹² by the HMBC correlation between the NH of Glu¹² to the C=O of Trp¹¹ and the ester bond between Glu¹² and Ser⁹ was determined by the HMBC correlation of the β -proton of Ser⁹ with the C=O of the Glu¹² at the δ position. The last residue of the sequence was deduced to be Glu¹³ and Phe¹⁴ by comparison of the structure with other known microviridins. Finally, composition of the macrolactam was identified by the HMBC correlation between the ϵ -NH of Lys⁶ and the C=O of the δ position of Glu¹³.

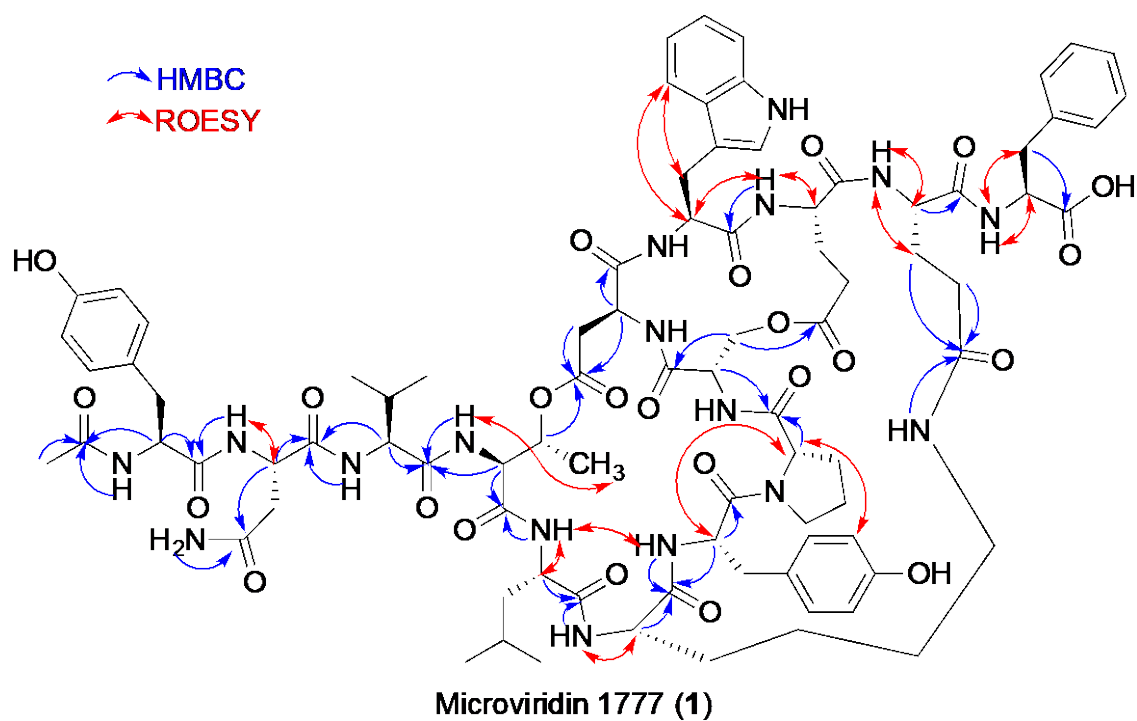


Figure 3. Structure of microviridin 1777 (1) isolated from *M. aeruginosa* EAWAG 127a with key HMBC (blue) and ROESY (red) correlations highlighted.

Table 1. NMR Data Analysis of Microviridin 1777 (1).

AA	pos.	δ_C , type	δ_H (J in Hz)	HMBC ^a
Ac	1	169.1, C		
	2	22.3, CH ₃	1.75, s	Ac(1)
Tyr ¹	1	171.6, C		
	2	54.3, CH	4.43, m	Tyr ¹ (1), Tyr ¹ (3), Tyr ¹ (4), Ac(1)
	3	36.7 CH ₂	2.89, m	Tyr ¹ (2), Tyr ¹ (4), Tyr ¹ (5,9)
			2.61, m	Tyr ¹ (2), Tyr ¹ (4), Tyr ¹ (5,9)
	4	127.9, C		
	5,9	129.9, CH	7.02, m	Tyr ¹ (3), Tyr ¹ (4), Tyr ¹ (6,8), Tyr ¹ (9)

	6,8	114.8, CH	6.63, m	Tyr ¹ (9)
	7	155.7, C		
	OH		9.18, s	
	NH		8.03, d (8.2)	Ac(1), Tyr ¹ (3), Tyr ¹ (3)
Asn ²	1	170.6, C		
	2	49.7, CH	4.61, m	Asn ² (1), Asn ² (3), Asn ² (4)
	3	36.6, CH ₂	2.46, m	Asn ² (1), Asn ² (3), Asn ² (4)
			2.59, m	Asn ² (1), Asn ² (3), Asn ² (4)
	4	171.5, C		
	NH		8.38, d (7.7)	Tyr ¹ (1), Asn ² (2), Asn ² (3)
	NH ₂		6.89, m	Asn ² (3), Asn ² (4)
Val ³	1	171.0, C		
	2	57.1, CH	4.40, m	Val ³ (1), Val ³ (3), Val ³ (4), Val ³ (4'), Asn ² (1)
	3	30.6, CH	2.01, m	Val ³ (1), Val ³ (2), Val ³ (4), Val ³ (4')
	4	18.9, CH ₃	0.84, d (6.8)	Val ³ (2), Val ³ (3), Val ³ (4')
	4'	17.4, CH ₃	0.80, d (6.8)	Val ³ (2), Val ³ (3), Val ³ (4)
	NH		7.53, d (8.9)	Asn ² (1), Val ³ (2)
Thr ⁴	1	168.7, C		
	2	54.6, CH	4.57, m	Thr ⁴ (1), Thr ⁴ (3)
	3	70.8, CH	5.33, m	Val ³ (1), Thr ⁴ (2), Thr ⁴ (4), Asp ¹⁰ (4)
	4	16.8, CH ₃	1.15, d (6.4)	Thr ⁴ (2), Thr ⁴ (3)

	NH		7.70, d (8.8)	Val ³ (1)
Leu ⁵	1	170.2, C		
	2	50.7, CH	4.23, m	Leu ⁵ (1), Leu ⁵ (3)
	3	40.2, CH ₂	1.37, m	Leu ⁵ (2), Leu ⁵ (4), Leu ⁵ (5), Leu ⁵ (5')
			1.56, m	Leu ⁵ (2), Leu ⁵ (4), Leu ⁵ (5), Leu ⁵ (5')
	4	23.9, CH	1.45, m	Leu ⁵ (3), Leu ⁵ (5), Leu ⁵ (5'),
	5	21.3, CH ₃	0.74, d (6.4)	Leu ⁵ (3), Leu ⁵ (4), Leu ⁵ (5')
	5'	23.1, CH ₃	0.82, d (6.6)	Leu ⁵ (3), Leu ⁵ (4), Leu ⁵ (5)
	NH		8.27, d (8.6)	Thr ⁴ (1), Leu ⁵ (2)
Lys ⁶	1	170.1, C		
	2	52.5, CH	4.12, m	Lys ⁶ (1), Lys ⁶ (3), Lys ⁶ (4)
	3	32.1, CH ₂	1.53, m	Lys ⁶ (2), Lys ⁶ (4)
	4	21.6, CH ₂	1.22, m	Lys ⁶ (3), Lys ⁶ (5), Lys ⁶ (6)
	5	28.7, CH ₂	1.44, m	Lys ⁶ (3), Lys ⁶ (4), Lys ⁶ (6)
			1.29, m	Lys ⁶ (3), Lys ⁶ (4), Lys ⁶ (6)
	6	37.7, CH ₂	2.94, m	Lys ⁶ (5), Lys ⁶ (6)
			3.14, m	Lys ⁶ (5), Lys ⁶ (6)
	αNH		6.70, d (7.4)	Leu ⁵ (1)
	εNH		6.91, m	Lys ⁶ (6), Glu ¹³ (5)
Tyr ⁷	1	170.9, C		
	2	51.6, CH	4.36, m	Tyr ⁷ (1), Tyr ⁷ (3), Tyr ⁷ (4), Lys ⁶ (1)
	3	37.4, CH ₂	2.77, m	Tyr ⁷ (1), Tyr ⁷ (2), Tyr ⁷ (4), Tyr ⁷ (5,9)

			2.67, m	Tyr ⁷ (1), Tyr ⁷ (2), Tyr ⁷ (4), Tyr ⁷ (5,9)
	4	126.3, C		
	5,9	130.0, CH	6.95, m	Tyr ⁷ (3), Tyr ⁷ (4), Tyr ⁷ (6,8) Tyr ⁷ (7)
	6,8	115.1, CH	6.67, m	Tyr ⁷ (4), Tyr ⁷ (5,9) Tyr ⁷ (7)
	7	156.2, C		
	OH		9.34, s	
	NH		8.33, d (8.3)	Lys ⁶ (1), Tyr ⁷ (2)
Pro ⁸	1	170.7, C		
	2	60.4, CH	3.47, m	Pro ⁸ (1), Pro ⁸ (3)
	3	30.7, CH ₂	1.62, m	
			1.53, m	
	4	31.8, CH ₂	1.92, m	
			2.09, m	
	5	46.0, CH ₂	3.41, m	
			3.23, m	
Ser ⁹	1	169.0, C		
	2	52.1, CH	4.47, m	Pro ⁸ (1)
	3	61.8, CH ₂	4.20, m	Ser ⁹ (1)
			4.54, m	Glu ¹² (5)
	NH		7.36, s	Asp ¹⁰ (1)
Asp ¹⁰	1	171.5, C		
	2	54.1, CH	4.42, m	Asp ¹⁰ (4)
	3	34.5, CH ₂	2.63, m	Asp ¹⁰ (1), Asp ¹⁰ (4)

			2.77, m	
	4	169.4, C		
	NH		n.o.	
Trp ¹¹	1	171.1, C		
	2	53.4, CH	4.58, m	Trp ¹¹ (1)
	3	25.8, CH ₂	3.18, m	
			3.28, m	Trp ¹¹ (1)
	1' NH		10.86, s	Trp ¹¹ (2'), Trp ¹¹ (3'), Trp ¹¹ (8'), Trp ¹¹ (9')
	2'	123.3, CH	7.19, m	Trp ¹¹ (3'), Trp ¹¹ (8'), Trp ¹¹ (9')
	3'	109.4, C		
	4'	118.0, CH	7.47, d (7.8)	Trp ¹¹ (3'), Trp ¹¹ (5'), Trp ¹¹ (8'), Trp ¹¹ (9')
	5'	121.0, CH	7.05, m	Trp ¹¹ (4'), Trp ¹¹ (6'), Trp ¹¹ (7'), Trp ¹¹ (8'), Trp ¹¹ (9')
	6'	118.4, CH	6.97, m	Trp ¹¹ (4'), Trp ¹¹ (5'), Trp ¹¹ (7'), Trp ¹¹ (8'), Trp ¹¹ (9')
	7'	111.4, CH	7.28, m	Trp ¹¹ (6'), Trp ¹¹ (8')
	8'	127.2, C		
	9'	136.1, C		
	αNH		n.o.	
Glu ¹²	1	n.o.		
	2	52.5, CH	4.16, m	
	3	24.4, CH ₂	1.61, m	
			1.85, m	

Glu ¹³	4	29.4, CH ₂	1.33, m	
			2.08, m	
	5	171.0, C		
	NH		6.50, d (8.3)	
	1	169.7, C		
	2	52.4, CH	4.03, m	
	3	27.9, CH ₂	2.08, m	
			1.53, m	
	4	31.8, CH ₂	1.91, m	Glu ¹³ (5)
	5	171.0, C		
Phe ¹⁴	NH		7.41, m	
	1	172.7, C		
	2	55.2, CH	4.02, m	
	3	37.4, CH ₂	3.07, dd	Phe ¹⁴ (1), Phe ¹⁴ (2), Phe ¹⁴ (4), Phe ¹⁴ (5,9)
			(13.5, 5.2)	
			2.91, m	Phe ¹⁴ (2), Phe ¹⁴ (4), Phe ¹⁴ (5,9), Phe ¹⁴ (6,8)
	4	139.1, C		
	5,9	129.5, CH	7.16, m	Phe ¹⁴ (3), Phe ¹⁴ (4), Phe ¹⁴ (6,8), Phe ¹⁴ (7)
	6,8	127.5, CH	7.17, m	Phe ¹⁴ (4), Phe ¹⁴ (5,9), Phe ¹⁴ (7)
	7	125.4, CH	7.10, m	Phe ¹⁴ (5,9), Phe ¹⁴ (6,8)
	NH		7.29, m	

^aHMBC correlations are given from proton(s) stated to the indicated carbon atom. Abbreviation: n.o. for not observed.

Having the characterized and purified natural product in hand, we next investigated biological activity. The potency of microviridin 1777 (**1**) was evaluated against *T. platyurus*, which resulted in an LD₅₀ value of 95 µM. The cytotoxicity of purified **1** was around five times greater than the extract, but remains roughly one order of magnitude weaker than other cyanobacterial toxins.²⁷⁻³⁷

To understand the mechanism of action of microviridin 1777 (**1**), an enzyme assay with trypsin, chymotrypsin, and elastase was performed (**Figure 4**) using modified literature procedures.^{31,45} Whereas only weak inhibition of trypsin could be observed (IC₅₀ > 10 µM), microviridin 1777 (**1**) displayed stronger activity against elastase with an IC₅₀ of 160 (± 10) nM. This is in line with data for other microviridins in the literature, where peptides carrying either leucine or valine at position 5 result in good activity against elastase.⁵⁰ Inhibitor **1** demonstrated more potent activity against chymotrypsin with an IC₅₀ value of 100 (± 10) nM, which to date represents the most potent chymotrypsin inhibitory activity of any microviridin (Supporting Information Table S4).

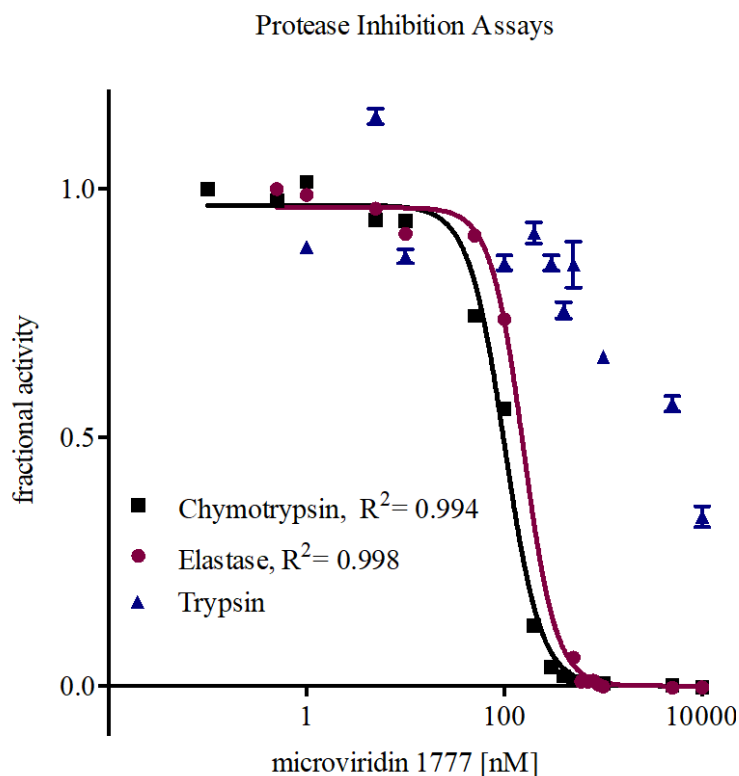


Figure 4. Fractional activity of **1** on the chymotrypsin, trypsin, and elastase protease inhibition assay (n=3).

A detailed analysis of the extracts by UHPLC-HRMS/HRMS revealed the presence of microviridin M (**2**),⁵¹ an N-terminally truncated version of microviridin 1777 (**1**) (**Figure 1**). This compound was produced heterologously from *E. coli* and its IC_{50} value against chymotrypsin and trypsin were higher than 34 μ M, and an IC_{50} value of 2.9 μ M was determined against elastase.⁵¹ This result highlights the importance of the microviridin 1777 (**1**) side chain for its serine protease activity. Additionally, microviridin O was not identified in the extract although the precursor peptide was found in the genome (EZJ55_03525).

Using the genomic approach, the NPRS gene cluster of the anabaenopeptin family of natural products was identified in the genome of *M. aeruginosa* EAWAG 127a using antiSMASH

(**Figure 5**). The cyclic peptides were expected to contain eight amino acids due to the presence of eight predicted adenylation domains. The organization of the cluster was in agreement with those previously isolated from *M. aeruginosa*^{61,62} and in *Planktothrix agardhii*.⁶³ The UHPLC-HRMS/HRMS data were analyzed with the GNPS platform using the DEREPLICATOR+ algorithm and three known anabaenopeptins (A, B and F) and oscillamide Y⁶⁴ were detected. The identities of the compounds was confirmed by their characteristic HRMS/HRMS pattern (Supporting Information, Figures S7-S10).⁶⁵ The molecular networking analysis distinguished two classes of anabaenopeptins which possessed either a tyrosine or arginine.

Continuing our genomic investigation, we localized the aeruginosin NRPS gene cluster using antiSMASH (**Figure 5**). As observed in previous analysis of *M. aeruginosa*,⁶¹ the gene responsible for the production of the halogenase (*aerJ*) was not identified in the genome of the EAWAG 127a strain. Analysis of the extract by UHPLC-HRMS/HRMS data led to the identification of aeruginosin 102 derivative **3** missing the sulfate group (**Figure 6**).^{66,67} The identity of the natural product was corroborated by tandem MS (Supporting Information, **Figure S11**).

Finally, the biosynthetic gene cluster responsible for the production of piricyclamide was predicted by antiSMASH with a similarity percentage of more than 60% to a reported gene cluster from this family. Interestingly, the putative piricyclamide gene cluster in *M. aeruginosa* EAWAG 127a does not possess the geranyltransferase *pirF* (**Figure 5**).⁷⁰ These modified cyclic peptides were not detected in the current study.

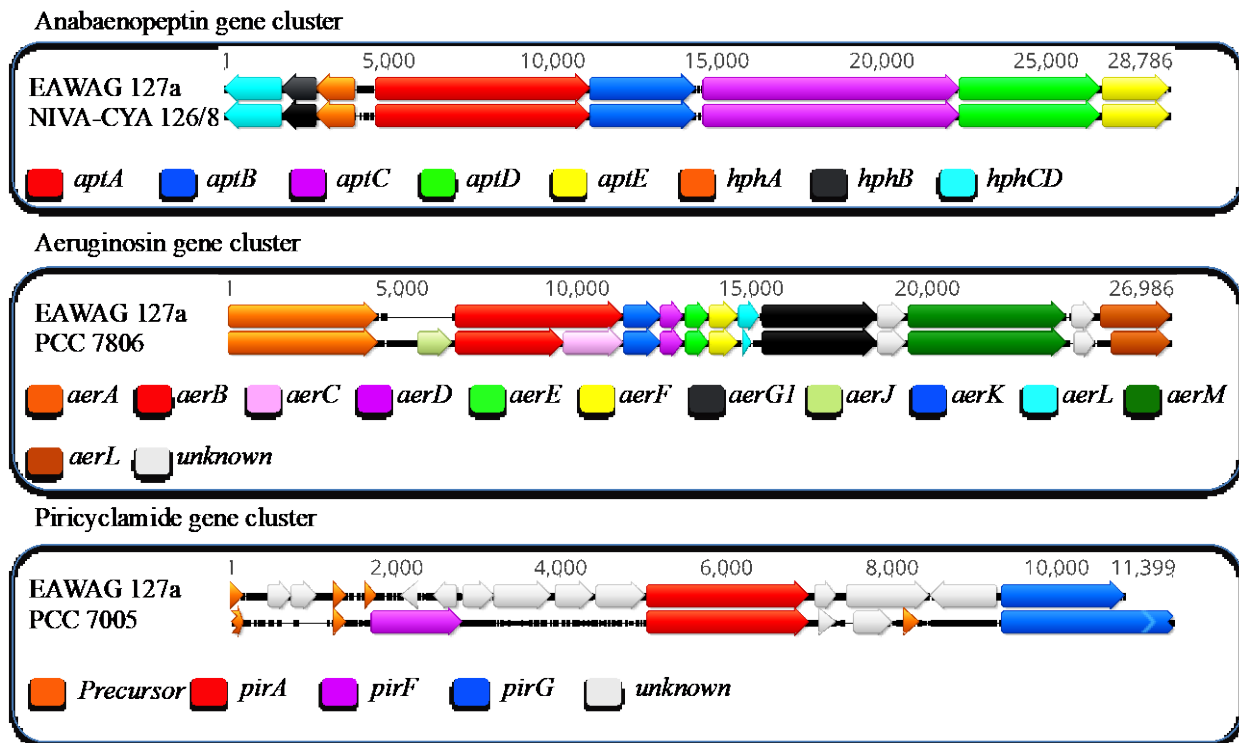


Figure 5. Alignment of anabaenopeptin gene cluster from *Planktothrix agardhii* NIVA-CYA 126/8 (accession ID: EF672686.1, from left to right: WP_042156037.1 to WP_042156027.1)⁶³ and the one of *M. aeruginosa* EAWAG 127a (accession ID: SRLN000000000, from left to right: EZJ55_13000 to EZJ55_13035). Alignment of aeruginosin gene cluster from *Microcystis aeruginosa* PCC 7806 (accession ID: AM778955.1, from left to right: CAO90437.1 to CAO90422.1)^{68,69} and the one of *M. aeruginosa* EAWAG_127a (accession ID: SRLN000000000, from left to right: EZJ55_15670 to EZJ55_15725). Alignment of piricyclamide gene cluster from *M. aeruginosa* PCC 7005 (accession ID: JQ951924.1, from left to right: AFK79987.1 to AFK79998.1)⁷¹ and the one of *M. aeruginosa* EAWAG 127a (accession ID: SRLN000000000, from left to right: EZJ55_18570 to EZJ55_18645).

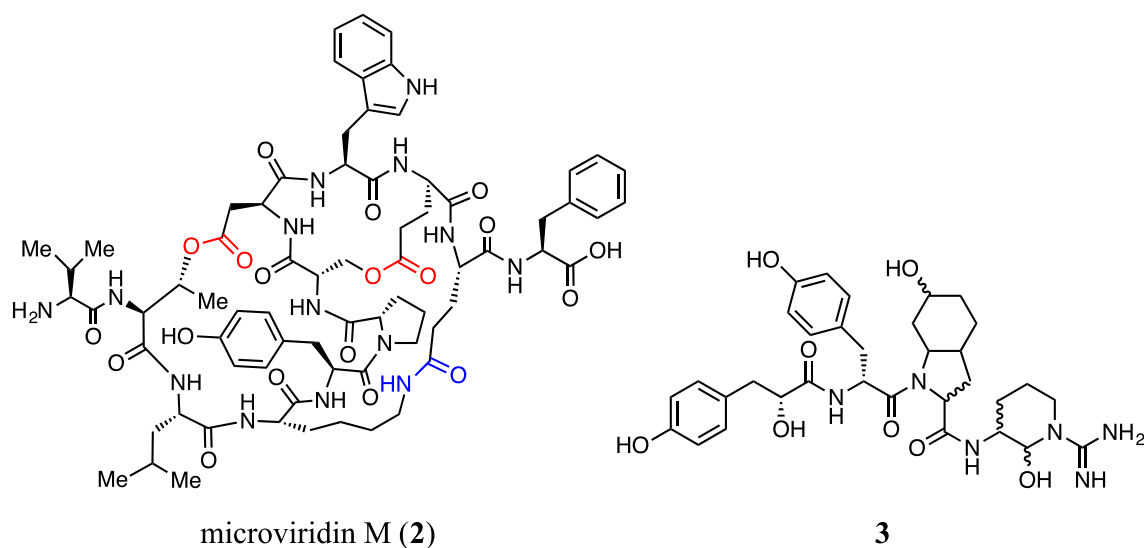


Figure 6. Structure and configurations of microviridin M (2) and aeruginosin 102 derivative **3** missing the sulfate group at the phenol position of the *p*-hydroxyphenyllactic acid residue

In conclusion, *M. aeruginosa* EAWAG 127a was selected as a promising producer of protease inhibitors during a preliminary screening against *T. platyurus*. To identify putative toxins produced by this strain, the secondary metabolite peptides were studied using a metabologenomic approach. The DNA of the cyanobacterium was isolated, sequenced, and analyzed using freely available online tools. Additionally, the extract was obtained, analyzed with UHPLC HRMS/HRMS, and the resulting data were uploaded on the GNPS platform. The combined strategy led to the identification and structure elucidation of microviridin 1777 (**1**), the most potent chymotrypsin inhibitor from the microviridin family. Additionally, inhibitor **1** was found to possess an IC_{50} value of 95 μ M against *T. platyurus*, which reinforces the ecological importance of cyclic peptides produced by cyanobacteria. Furthermore, anabaenopeptins A, B and F, oscillamide Y, the aeruginosin 102 derivative **3**, and a truncated version of microviridin 1777 (**1**), microviridin M (**2**), were identified in the extract using the GNPS platform. These results highlight the efficacy and robustness of database mining tools. Finally, assessment of freshwater quality should include other compound classes beyond microcystin.

Experimental Section

General Experimental Procedures. All ^1H and ^{13}C NMR spectra were recorded using a Bruker Avance 600 MHz (^1H) and 151 MHz (^{13}C) spectrometer at RT and 310 K. Chemical shifts (δ -values) are reported in ppm, spectra were calibrated relative to the residual proton chemical shifts (DMSO- d_6 , $\delta = 2.50$) and the residual carbon chemical shifts (DMSO- d_6 , $\delta = 39.52$) of the solvents, multiplicity is reported as follows: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet or unresolved and coupling constant J in Hz. The Mass Spectrometry Service of the University of Zurich recorded the HRMS-ESI and MS/MS spectra on a QExactive instrument (Thermo Fisher Scientific, Bremen, Germany) equipped with BEH C₁₈ 1.7 μM 2.1 \times 150 mm column, a heated electrospray (ESI) ionization source and connected to a Dionex Ultimate 3000 UHPLC system. Preparative HPLC was carried out on a Shimadzu HPLC system equipped with a Gemini column (5 μm , NX-C₁₈ 110Å, 250 \times 21.2 mm) and fraction collector. Fraction analysis was conducted on a Thermo Fisher Scientific Ultimate 3000 LC instrument coupled to a triple quadrupole Quantum Ultra EMR MS (H-ESI II ion source) equipped with a reversed-phase column (Kinetex[®] EVO C₁₈, 1.7 μm , 100 Å, 50 \times 2.1 mm Phenomenex). The solvents were purchased from Sigma-Aldrich, Acros or Fluka in HPLC grade. Removing the solvent under reduced pressure was carried out with a rotary evaporator with a water bath temperature of 40 °C. A Synergy H1 microplate reader from BioTek was used to record fluorescence and absorbance of the enzyme assays. The enzymes trypsin type I from bovine pancreas, α -chymotrypsin II from bovine pancreas and elastase pancreatic type I from porcine as well as the substrate SucALA₃-pNA were purchased from Sigma Aldrich. The substrates Glutaryl-Gly-Arg-AMC and Glutaryl-Gly-Gly-Phe-AMC were purchased from Bachem AG. Cryst Freeze dryer ALPHA 1-4 LD plus from Kuhner AG was used to perform lyophilization.

Sources and Cultivation of Cyanobacteria. The strain *M. aeruginosa* EAWAG 127a is part of the cyanobacteria collection of the University of Zurich, formerly located at EAWAG (Swiss Federal Institute of Aquatic Science and Technology). The strain was grown in a 60 L batch reactor (Z medium) with a light/dark cycle of 12:12 h and continuous airflow. The biomass was obtained by centrifugation ($4500 \times g$, 30 min, 4 °C) of 1 L of the culture and removing the medium by filtration.

Extraction and Isolation. The wet biomass (10.22 g) was extracted with an aq. MeOH soln. (60%, 150 mL), and the resulting mixture was sonicated (3×3 min with 5 min breaks, to avoid overheating), harvested by centrifugation ($9000 \times g$, 30 min, 4 °C) and the soln. was carefully decanted. This extraction procedure was repeated once. The extracted MeOH solutions were combined, evaporated under reduced pressure and the residue was lyophilized. The dry crude material (251.9 mg) was dissolved in an aq. MeOH soln. (40%), transferred in a Falcon tube, harvested by centrifugation ($9000 \times g$, 10 min, 4 °C). The mixture was loaded on a C₁₈ solid-phase extraction (SPE) column (Supelco Discovery DSC-18, 2 g, that was pre-washed with MeOH and H₂O and conditioned with 40% aq. MeOH soln.). The loaded column was rinsed with aq. MeOH soln. (40%) followed by extensive elution with aq. MeOH soln. (60%). The remaining solid in the Falcon tube was dissolved in MeOH (60%), harvested by centrifugation ($9000 \times g$, 10 min, 4 °C) and also loaded on the C₁₈ SPE column and the compounds were eluted with aq. MeOH soln (60%). After this pre-separation, the 60% MeOH fractions were combined and the MeOH was removed under reduced pressure and the mixture was lyophilized. The greenish powder was resolubilized in MeOH (60%), filtered through a HPLC filter and the compound was isolated by preparative HPLC (Gemini, 5 μ m, NX-C₁₈ 110 Å, 250 \times 21.2 mm, 20 mL/min). The MeCN/H₂O gradient started with 5% to 40% in 5 min, then from 40% to 41.7% in 12 min and finally from

41.7% to 100% in 1 min. The column was re-equilibrated for 5 min to the initial gradient of 5%. Microviridin 1777 (**1**, 11.42 min) was isolated as a colorless solid (0.84 mg/L). Oscillamide Y (11.78 min) was isolated as a colorless solid (0.45 mg/L).

Microviridin (1): colorless solid; ^1H NMR (600 MHz, DMSO- d_6) δ 10.86 (s, 1H), 9.34 (s, 1H), 9.18 (s, 1H), 8.38 (d, $J = 7.7$ Hz, 1H), 8.33 (d, $J = 8.3$ Hz, 1H), 8.27 (d, $J = 8.6$ Hz, 1H), 8.03 (d, $J = 8.2$ Hz, 1H), 7.70 (d, $J = 8.8$ Hz, 1H), 7.53 (d, $J = 8.9$ Hz, 1H), 7.47 (d, $J = 7.8$ Hz, 1H), 7.42 – 7.38 (m, 1H), 7.36 (s, 1H), 7.29 – 7.27 (m, 2H), 7.21 – 7.13 (m, 5H), 7.13 – 7.07 (m, 1H), 7.07 – 7.00 (m, 3H), 6.99 – 6.92 (m, 3H), 6.92 – 6.85 (m, 2H), 6.70 (d, $J = 7.4$ Hz, 1H), 6.69 – 6.65 (m, 3H), 6.66 – 6.61 (m, 2H), 6.50 (d, $J = 8.3$ Hz, 1H), 5.37 – 5.30 (m, 1H), 4.63 – 4.52 (m, 4H), 4.50 – 4.45 (m, 1H), 4.45 – 4.32 (m, 4H), 4.26 – 4.18 (m, 2H), 4.18 – 4.09 (m, 2H), 4.07 – 3.98 (m, 2H), 3.50 – 3.45 (m, 1H), 3.43 – 3.38 (m, 1H), 3.28 – 3.21 (m, 2H) 3.21 – 3.11 (m, 2H), 3.07 (dd, $J = 13.5, 5.2$ Hz, 1H), 2.96 – 2.86 (m, 3H), 2.81 – 2.71 (m, 2H), 2.70 – 2.56 (m, 4H), 2.53 – 2.44 (m, 2H), 2.12 – 1.99 (m, 4H), 1.95 – 1.87 (m, 1H), 1.86 – 1.82 (m, 1H), 1.75 (s, 3H), 1.67 – 1.58 (m, 3H), 1.58 – 1.49 (m, 3H), 1.48 – 1.40 (m, 2H), 1.40 – 1.27 (m, 3H), 1.26 – 1.19 (m, 2H), 1.15 (d, $J = 6.4$ Hz, 3H), 0.85 – 0.84 (m, 3H), 0.84 (d, $J = 6.8$ Hz, 3H), 0.82 (d, $J = 6.6$ Hz, 3H), 0.80 (d, $J = 6.8$ Hz, 3H), 0.75 (d, $J = 6.4$ Hz, 3H); ^{13}C NMR (151 MHz, DMSO- d_6) δ 172.75, 171.63, 171.55, 171.53, 171.09, 171.06, 170.98, 170.90, 170.75, 170.64, 170.23, 170.05, 169.73, 169.39, 169.13, 168.97, 168.75, 156.19, 155.69, 139.07, 136.15, 129.96, 129.87, 129.48, 127.95, 127.50, 127.15, 126.31, 125.43, 123.36, 120.99, 118.39, 118.02, 115.07, 114.78, 111.43, 109.43, 70.83, 61.80, 60.44, 57.05, 55.19, 54.60, 54.27, 54.11, 53.37, 52.47, 52.43, 52.35, 52.07, 51.55, 50.70, 49.71, 46.03, 40.18, 37.66, 37.36, 36.66, 36.64, 34.54, 32.11, 31.82, 30.68, 30.65, 29.43, 28.65, 27.92, 25.75, 24.38, 23.87, 23.10, 22.32, 21.60, 21.25, 18.93, 17.40, 16.81; HRESI m/z 889.9058 Da $[\text{M} + 2\text{H}]^{2+}$ (calc. for $\text{C}_{87}\text{H}_{113}\text{N}_{17}\text{O}_{24}$ 889.9067 Da).

Mass spectrometry and molecular networking analysis. The wet biomass (4.71 g) was extracted with an aq. MeOH soln. (60%, 150 mL), and the resulting mixture was sonicated (3 × 3 min with 5 min breaks to avoid overheating), harvested by centrifugation (9000 × g, 30 min, 4 °C) and the mixture was carefully decanted. This extraction procedure was repeated once. The MeOH solutions were combined, evaporated under reduced pressure, and the residue was lyophilized. The dry crude material (66 mg) was dissolved in aq. MeOH soln. (80%) with a concentration of 8 mg/mL. The solutions were harvested by centrifugation in Eppendorf tubes. UHPLC-DAD-HRMS/HRMS spectra were measured in positive mode on a BEH C₁₈ 2 × 150 mm, with a normalized collision energy (NCE) of 15, 22 and 30. The MeCN/H₂O gradient varied from 3% to 100% in 12 min and the column was washed for 3 min. The column was recalibrated in 2 min to the initial gradient of 3%. The UHPLC-HRMS/HRMS data were uploaded on the GNPS database⁴ and the analyses were performed with the DEREPLICATOR and DEREPLICATOR+ for the dereplication process. A molecular network was created using the online workflow at GNPS. The data was filtered by removing all MS/MS peaks within +/- 17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 peaks in the +/- 50 Da window throughout the spectrum. The data was then clustered with MS-Cluster with a parent mass tolerance of 1.0 Da and a MS/MS fragment ion tolerance of 0.5 Da to create consensus spectra. Further, consensus spectra that contained less than two spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.6 and more than 6 matched peaks. Further edges between two nodes were kept in the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were then searched against GNPS spectral libraries. The library spectra were filtered in the same manner as the input data. All matches kept between network spectra and library spectra were required to have a score above 0.7

and at least 6 matched peaks. Analog search was enabled against the library with a maximum mass shift of 100.0 Da.

Trypsin and chymotrypsin assay. The trypsin and chymotrypsin inhibition assays were carried out using the procedure from Gademann et al.³¹ The fluorescence measurements were carried out using a fluorescence microplate reader set at $\lambda_{\text{ex}} = 355$ nm and $\lambda_{\text{em}} = 480$ nm and measuring the kinetic data points every 30 s for 1 h in a 96 dark well plate. Trypsin and chymotrypsin were dissolved in a Tris-HCl buffer (50 mM, pH = 7.5) containing NaCl (100 mM), MgCl₂ (10 mM), CaCl₂ (1 mM), BSA (0.1%) Triton X-100 (0.01%). To ensure the solubility of the substrate, DMSO (7.5%) was added to the buffer and this buffer was also used as the blank. The final protease and substrate concentrations are listed in a table reported in the supporting information (**Table S1**). The final inhibitor concentrations ranged between 0.1 and 10 nM. The inhibition assay was carried out by incubation of protease solution (50 μ L) with microviridin 1777 (**1**) solution (50 μ L) for 30 min at RT. After addition of the substrate solution (50 μ L) using a multichannel pipette dispenser, the data were immediately collected on the microplate reader. The control experiment was performed by the incubation of blank buffer (50 μ L) and the enzyme solution (50 μ L) for 30 min following by the addition of the substrate solution (50 μ L). The inhibition curves and the 50% inhibition values (IC₅₀) were calculated using the linear regression of the substrate/enzyme rate respectively substrate/inhibitor rate of the first 10 min and then dividing the inhibition slope through the substrate slope. To determine the fractional activity and the IC₅₀ value a nonlinear regression sigmoidal four-parameter curve was determined using Prism 7.

Elastase assay. The elastase inhibition assay was carried out using a modified procedure recently published by Dittmann and co-workers.⁴⁵ Elastase, its substrate and microviridin were dissolved in a Tris-HCl buffer (50 mM, pH = 7.5) and the enzyme and the inhibitor were incubated

for 30 min at RT. After the substrate addition, the absorbance was immediately collected on a microplate reader at 410 nm over 15 min at 37 °C. Linear regression was performed of the substrate/enzyme rate respectively substrate inhibitor rate on the first 5 min and then dividing the inhibition slope through the substrate slope. To determine the fractional activity and the IC₅₀ value a nonlinear regression sigmoidal four-parameter curve was determined using Prism 7.

Toxicity assay. The toxicity assay was conducted with Thamnotoxkit F, crustacean toxicity screening test for freshwater, and it was purchased from MicobioTest Inc., Kleimoer 15, 9030 Mariakerke (Gent), Belgium. Throughout the assay the standard operational procedure was followed. The standard freshwater was prepared, adding the concentrated salt solutions in the given sequence to 800 mL of MilliQ water in a 1 L volumetric flask, followed by the addition of water to the 1 L mark and bubbling air with an aquarium pump for 20 min. The cyst hatching was performed by a 1:8 dilution of the standard freshwater with MilliQ water (2.5 mL:17.5 mL) to prepare the hatch medium. The hatch medium was added to the tube with cysts. During a 30 min period the tube with the cysts was shaken at regular intervals. The cysts were then transferred to a petri dish with 15 mL of diluted standard freshwater. The petri dish was shaken gently to equally distribute the cysts and then incubated for 22 h at 25 °C under a continuous illumination of 3600 Lux. Between 10 and 20 shrimps were transferred to each of the 24 wells and the different concentrations of microviridin 1777 (**1**) and extract as well as a positive and negative control were performed in triplicates and the results are found in the supporting information (Tables S2 and S3). To the extract soln. DMSO (0.2%) was added to enhance the solubility.

DNA isolation and sequencing. The DNA of *M. aeruginosa* EAWAG 127a was obtained using a modified protocol.⁴³ The cyanobacterium *M. aeruginosa* EAWAG 127a culture (10 mL) was transferred into a Falcon tube, harvested by centrifugation (4,500 × g, 30 min, 4 °C) and the

supernatant was discarded to obtain the biomass. The biomass was transferred in an Eppendorf tube using the buffer T1/10 (500 μ L, 10 mM, Tris-HCl, 0.1 mM EDTA- Na_2 , pH = 8). A mixture of phenol:chloroform (450 μ L, 1:1 v/v), sterile glass beads (150 μ L) and an aq. SDS solution (20 μ L, 10%) were added, vortexed (1 min), cooled on ice (1 min) and this process was repeated three times. The mixture was centrifuged ($16,100 \times g$, 15 min, RT) and the supernatant was transferred to an Eppendorf tube and the aq. solution was washed with phenol (1. eq. volume, 450 μ L), phenol:chloroform (1. eq. volume, 450 μ L, 1:1 v/v, 3x) and with chloroform (1. eq. volume, 450 μ L, 3x). An aq. solution ethanol (2.5 eq. volume, 1.125 mL) and an aq. potassium acetate solution (0.1 eq. volume, 45 μ L, 3 M, pH = 5.2) were added and left at -20°C overnight to let the DNA precipitate. The precipitated DNA was taken out, washed with an aq. ethanol solution (70%) and dried under a gentle nitrogen flow.^{72,73} The resulting DNA was dissolved in water (100 μ L) and the concentration was analyzed by nanodrop (438.5 ng/ μ L, 43.8 μ g).

Library Preparation and Sequencing. The SMRTbell library was produced using the DNA Template Prep Kit 1.0 (Pacific Biosciences p/n 100-259-100). The concentration of the input genomic DNA was measured using a Qubit Fluorometer dsDNA Broad Range assay (Life Technologies p/n 32850). gDNA (11 μ g) were mechanically sheared to an average size distribution of 15-20 kb using a Covaris g-Tube (Kbiosciences p/n 520079). The resulting size distribution was analyzed on a Femto Pulse automated electrophoresis system (Agilent). Sheared gDNA (6 μ g) was DNA damage- and end-repaired using polishing enzymes. A blunt end ligation reaction followed by exonuclease treatment was performed to create the SMRTbell template. Size-selection of the SMRTbell library was done on a Blue Pippin device (Sage Science) to enrich for fragments > 11 kb. The sized-selected library was quality inspected and quantified on a Femto Pulse (Agilent) and on a Qubit Fluorometer (Life Technologies) respectively. A ready to sequence SMRTbell-

Polymerase Complex was created using the P6 DNA/Polymerase binding kit 2.0 (Pacific Biosciences p/n 100-236-500) according to the manufacturer's instructions. The Pacific Biosciences RS2 instrument was programmed to load and sequence the library on one SMRT cell v3.0 (Pacific Biosciences p/n100-171-800), taking a movie of 360 minutes per cell. MagBead loading (Pacific Biosciences p/n 100-133-600) was chosen in order to improve the enrichment of longer fragments. After the run a sequencing report was generated via the SMRT portal in order to assess the adapter dimer contamination, the sample loading efficiency, the obtained average read-length and the number of filtered sub-reads.

DNA assembly. PacBio raw data was assembled using HGAP3 in SMRT Analysis 2.3.0 with default parameters. Assembled contig sequences were imported into SMRT Link 6.0, and were further polished using ARROW with PacBio raw data. Polishing consensus sequences were annotated using PROKKA (v1.0).⁷⁴

Genomic analysis. The genomic analysis of *M. aeruginosa* EAWAG 127a was performed using the online tool antiSMASH (5.0.0)⁶⁻⁹ and RODEO,¹⁵⁻¹⁶ and the data were further analyzed using the software Geneious (10.2.5). This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession SRLN00000000. The version described in this paper is version SRLN01000000. The annotation of the genes was partially achieved using the NCBI services. The gene cluster alignments used the following information: microviridin B gene cluster from *M. aeruginosa* NIES 843, accession number AM943877.1,⁴⁴ anabaenopeptin gene cluster from *Planktothrix agardhii* NIVA-CYA 126/8, accession number EF672686.1,⁶³ aeruginosin gene cluster from *M. aeruginosa* PCC 7806, accession number AM778955.1^{68,69} and piricyclamide gene cluster from *M. aeruginosa* PCC 7005, accession number JQ951924.1.⁷¹

ASSOCIATED CONTENT

Supporting Information.

The following files are available free of charge.

HPLC chromatogram, MS, UV and NMR spectra, molecular networking analysis, serine protease and toxicity assays, and lists of microviridin natural products, detected RiPPs, and peptides can be found in the supporting information. (PDF)

AUTHOR INFORMATION

Corresponding Author

karl.gademann@uzh.ch

ORCID

Simon Sieber: 0000-0003-1763-2525

Douglas A. Mitchell: 0000-0002-9564-0953

Karl Gademann: 0000-0003-3053-0689

Note

The authors declare no competing financial interests.

The UHPLC-HRMS/HRMS data of *M. aeruginosa* EAWAG 127a were uploaded on the GNPS platform and can be accessed using the dataset name MSV000083491. This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession SRLN00000000.

The version described in this paper is version SRLN01000000.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

ACKNOWLEDGMENTS

We are grateful for the help of Prof. Dr. L. Bigler from the mass spectrometry laboratory, S. Grüter and Dr. W. Qi from the Functional Genomics Center Zurich, Dr. E. Piel for skillful technical support, and S. Ramesh for preliminary bioinformatics analysis.

Funding Sources

The research was funded in part by the Swiss National Science Foundation (CRSII5_186410 and 200021_182043) and the US National Institutes of Health (GM123998 to D.A.M.).

REFERENCES

- (1) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2012**, 75, 311–335.
- (2) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2016**, 79, 629–661.
- (3) Nguyen, D. D.; Wu, C.-H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; Jackson, C.; Ballesteros, J.; Sanchez, J.; Watrous, J. D.; Phelan, V. V.; van de Wiel, C.; Kersten, R. D.; Mehnaz, S.; De Mot, R.; Shank, E. A.; Charusanti, P.; Nagarajan, H.; Duggan, B. M.; Moore, B. S.; Bandeira, N.; Palsson, B. Ø.; Pogliano, K.; Gutiérrez, M.; Dorrestein, P. C. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, 110, E2611–E2620.
- (4) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapon, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W.-T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.;

Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C.-C.; Floros, D. J.; Gavilan, R. G.; Kleigrew, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C.-C.; Yang, Y.-L.; Humpf, H.-U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya P., C. A.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P.-M.; Phapale, P.; Nothias, L.-F.; Alexandrov, T.; Litaudon, M.; Wolfender, J.-L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D.-T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B. Ø.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. *Nat. Biotechnol.* **2016**, *34*, 828–837.

- (5) Yang, J. Y.; Sanchez, L. M.; Rath, C. M.; Liu, X.; Boudreau, P. D.; Bruns, N.; Glukhov, E.; Wodtke, A.; de Felicio, R.; Fenner, A.; Wong, W. R.; Linington, R. G.; Zhang, L.; Debonsi, H. M.; Gerwick, W. H.; Dorrestein, P. C. *J. Nat. Prod.* **2013**, *76*, 1686–1699.
- (6) Medema, M. H.; Blin, K.; Cimermancic, P.; de Jager, V.; Zakrzewski, P.; Fischbach, M.

- A.; Weber, T.; Takano, E.; Breitling, R. *Nucleic Acids Res.* **2011**, *39*, W339–W346.
- (7) Blin, K.; Medema, M. H.; Kazempour, D.; Fischbach, M. A.; Breitling, R.; Takano, E.; Weber, T. *Nucleic Acids Res.* **2013**, *41*, W204–W212.
- (8) Weber, T.; Blin, K.; Duddela, S.; Krug, D.; Kim, H. U.; Bruccoleri, R.; Lee, S. Y.; Fischbach, M. A.; Müller, R.; Wohlleben, W.; Breitling, R.; Takano, E.; Medema, M. H. *Nucleic Acids Res.* **2015**, *43*, W237–W243.
- (9) Blin, K.; Wolf, T.; Chevrette, M. G.; Lu, X.; Schwalen, C. J.; Kautsar, S. A.; Suarez Duran, H. G.; de Los Santos, E. L. C.; Kim, H. U.; Nave, M.; Dickschat, J. S.; Mitchell, D. A.; Shelest, E.; Breitling, R.; Takano, E.; Lee, S. Y.; Weber, T.; Medema, M. H. *Nucleic Acids Res.* **2017**, *45*, W36–W41.
- (10) Skinnider, M. A.; Dejong, C. A.; Rees, P. N.; Johnston, C. W.; Li, H.; Webster, A. L. H.; Wyatt, M. A.; Magarvey, N. A. *Nucleic Acids Res.* **2015**, *43*, 9645–9662.
- (11) de Jong, A.; van Hijum, S. A. F. T.; Bijlsma, J. J. E.; Kok, J.; Kuipers, O. P. *Nucleic Acids Res.* **2006**, *34*, W273–W279.
- (12) de Jong, A.; van Heel, A. J.; Kok, J.; Kuipers, O. P. *Nucleic Acids Res.* **2010**, *38*, W647–W651.
- (13) van Heel, A. J.; de Jong, A.; Montalbán-López, M.; Kok, J.; Kuipers, O. P. *Nucleic Acids Res.* **2013**, *41*, W448–W453.
- (14) Tietz, J. I.; Schwalen, C. J.; Patel, P. S.; Maxson, T.; Blair, P. M.; Tai, H.-C.; Zakai, U. I.; Mitchell, D. A. *Nat. Chem. Biol.* **2017**, *13*, 470–478.
- (15) Schwalen, C. J.; Hudson, G. A.; Kille, B.; Mitchell, D. A. *J. Am. Chem. Soc.* **2018**, *140*, 9494–9501.
- (16) Hudson, G. A.; Burkhart, B. J.; DiCaprio, A. J.; Schwalen, C. J.; Kille, B.; Pogorelov, T.

- V.; Mitchell, D. A. *J. Am. Chem. Soc.* **2019**, *141*, 8228–8238.
- (17) Gurevich, A.; Mikheenko, A.; Shlemov, A.; Korobeynikov, A.; Mohimani, H.; Pevzner, P. A. *Nat. Microbiol.* **2018**, *3*, 319–327.
- (18) Mohimani, H.; Gurevich, A.; Shlemov, A.; Mikheenko, A.; Korobeynikov, A.; Cao, L.; Shcherbin, E.; Nothias, L.-F.; Dorrestein, P. C.; Pevzner, P. A. *Nat. Commun.* **2018**, *9*, 4035.
- (19) Nothias, L.-F.; Nothias-Esposito, M.; da Silva, R.; Wang, M.; Protsyuk, I.; Zhang, Z.; Sarvepalli, A.; Leyssen, P.; Touboul, D.; Costa, J.; Paolini, J.; Alexandrov, T.; Litaudon, M.; Dorrestein, P. C. *J. Nat. Prod.* **2018**, *81*, 758–767.
- (20) Schrimpe-Rutledge, A. C.; Sherrod, S. D.; McLean, J. A. *Curr. Opin. Chem. Biol.* **2018**, *42*, 160–166.
- (21) Ong, J. F. M.; Goh, H. C.; Lim, S. C.; Pang, L. M.; Chin, J. S. F.; Tan, K. S.; Liang, Z.-X.; Yang, L.; Glukhov, E.; Gerwick, W. H.; Tan, L. T. *Mar. Drugs* **2019**, *17*, 72.
- (22) Amiri Moghaddam, J.; Crüsemann, M.; Alanjary, M.; Harms, H.; Dávila-Céspedes, A.; Blom, J.; Poehlein, A.; Ziemert, N.; König, G. M.; Schäberle, T. F. *Sci. Rep.* **2018**, *8*, 16600.
- (23) Hou, X.-M.; Li, Y.-Y.; Shi, Y.-W.; Fang, Y.-W.; Chao, R.; Gu, Y.-C.; Wang, C.-Y.; Shao, C.-L. *J. Org. Chem.* **2019**, *84*, 1228–1237.
- (24) Nguyen, D. D.; Melnik, A. V.; Koyama, N.; Lu, X.; Schorn, M.; Fang, J.; Aguinaldo, K.; Lincecum Jr, T. L.; Ghequire, M. G. K.; Carrión, V. J.; Cheng, T. L.; Duggan, B. M.; Malone, J. G.; Mauchline, T. H.; Sanchez, L. M.; Kilpatrick, A. M.; Raaijmakers, J. M.; De Mot, R.; Moore, B. S.; Medema, M. H.; Dorrestein, P. C. *Nat. Microbiol.* **2017**, *2*, 16197.

- (25) Teta, R.; Della Sala, G.; Glukhov, E.; Gerwick, L.; Gerwick, W. H.; Mangoni, A.; Costantino, V. *Environ. Sci. Technol.* **2015**, *49*, 14301–14310.
- (26) Briand, E.; Bormans, M.; Gugger, M.; Dorrestein, P. C.; Gerwick, W. H. *Environ. Microbiol.* **2016**, *18*, 384–400.
- (27) Kleigrew, K.; Almaliti, J.; Tian, I. Y.; Kinnel, R. B.; Korobeynikov, A.; Monroe, E. A.; Duggan, B. M.; Di Marzo, V.; Sherman, D. H.; Dorrestein, P. C.; Gerwick, L.; Gerwick, W. H. *J. Nat. Prod.* **2015**, *78*, 1671–1682.
- (28) Ding, C. Y. G.; Pang, L. M.; Liang, Z.-X.; Goh, K. K. K.; Glukhov, E.; Gerwick, W. H.; Tan, L. T. *Mar. Drugs* **2018**, *16*, 505.
- (29) Le Manach, S.; Duval, C.; Marie, A.; Djediat, C.; Catherine, A.; Edery, M.; Bernard, C.; Marie, B. *Front. Microbiol.* **2019**, *10*, 1929.
- (30) Portmann, C.; Blom, J. F.; Gademann, K.; Jüttner, F. *J. Nat. Prod.* **2008**, *71* (7), 1193–1196.
- (31) Gademann, K.; Portmann, C.; Blom, J. F.; Zeder, M.; Jüttner, F. *J. Nat. Prod.* **2010**, *73*, 980–984.
- (32) Kohler, E.; Grundler, V.; Häussinger, D.; Kurmayer, R.; Gademann, K.; Pernthaler, J.; Blom, J. F. *Harmful Algae* **2014**, *39*, 154–160.
- (33) Scherer, M.; Bezold, D.; Gademann, K. *Angew. Chem., Int. Ed. Engl.* **2016**, *55*, 9427–9431.
- (34) Scherer, M.; Gademann, K. *Org. Lett.* **2017**, *19*, 3915–3918.
- (35) Portmann, C.; Blom, J. F.; Kaiser, M.; Brun, R.; Jüttner, F.; Gademann, K. *J. Nat. Prod.* **2008**, *71*, 1891–1896.
- (36) Portmann, C.; Sieber, S.; Wirthensohn, S.; Blom, J. F.; Da Silva, L.; Baudat, E.; Kaiser,

- M.; Brun, R.; Gademann, K. *J. Nat. Prod.* **2014**, *77*, 557–562.
- (37) Rohrlack, T.; Christoffersen, K.; Hansen, P. E.; Zhang, W.; Czarnecki, O.; Henning, M.; Fastner, J.; Erhard, M.; Neilan, B. A.; Kaebernick, M. *J. Chem. Ecol.* **2003**, *29*, 1757–1770.
- (38) Rohrlack, T.; Christoffersen, K.; Kaebernick, M.; Neilan, B. A. *Appl. Environ. Microbiol.* **2004**, *70*, 5047–5050.
- (39) Schwarzenberger, A.; Sadler, T.; von Elert, E. *J. Exp. Biol.* **2013**, *216*, 3649–3655.
- (40) Agrawal, M. K.; Zitt, A.; Bagchi, D.; Weckesser, J.; Bagchi, S. N.; von Elert, E. *Environ. Toxicol.* **2005**, *20*, 314–322.
- (41) Schwarzenberger, A.; Zitt, A.; Kroth, P.; Mueller, S.; von Elert, E. *BMC Physiol.* **2010**, *10*, 6.
- (42) Brahmi, M. M.; Portmann, C.; D'Ambrosio, D.; Woods, T. M.; Banfi, D.; Reichenbach, P.; Da Silva, L.; Baudat, E.; Turcatti, G.; Lingner, J.; Gademann, K. *Chem. - Eur. J.* **2013**, *19*, 4596–4601.
- (43) L. C. Guerrero, *PhD thesis*, Sevilla, **2014**, “Acción de los reguladores de la diferenciación en el establecimiento del patrón de heterocistos en la cianobacteria *Anabaena* sp PCC7120”.
- (44) Ziemert, N.; Ishida, K.; Liaimer, A.; Hertweck, C.; Dittmann, E. *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 7756–7759.
- (45) Weiz, A. R.; Ishida, K.; Quitterer, F.; Meyer, S.; Kehr, J.-C.; Müller, K. M.; Groll, M.; Hertweck, C.; Dittmann, E. *Angew. Chem., Int. Ed. Engl.* **2014**, *53*, 3735–3738.
- (46) Philmus, B.; Christiansen, G.; Yoshida, W. Y.; Hemscheidt, T. K. *ChemBioChem* **2008**, *9*, 3066–3073.

- (47) Philmus, B.; Guerrette, J. P.; Hemscheidt, T. K. *ACS Chem. Biol.* **2009**, *4*, 429–434.
- (48) Welker, M.; Brunke, M.; Preussel, K.; Lippert, I.; von Döhren, H. *Microbiology* **2004**, *150*, 1785–1796.
- (49) Welker, M.; von Döhren, H. *FEMS Microbiol. Rev.* **2006**, *30*, 530–563.
- (50) Ziemert, N.; Ishida, K.; Weiz, A.; Hertweck, C.; Dittmann, E. *Appl. Environ. Microbiol.* **2010**, *76*, 3568–3574.
- (51) (a) Gatte Picchi, D.; Weiz, A.; Ishida, K.; Hertweck, C.; Dittmann, E. *Appl. Environ. Microbiol.* **2014**, *80*, 1380–1387. (b) Ahmed, M.N.; Reyna-González, E.; Schmid, B.; Wiebach, V.; Süßmuth, R. D.; Dittmann, E.; Fewer, D. P. *ACS Chemical Biology* **2017**, *12*, 1538–1546.
- (52) Micallef, M. L.; D'Agostino, P. M.; Sharma, D.; Viswanathan, R.; Moffitt, M. C. *BMC Genomics* **2015**, *16*, 31.
- (53) Ishitsuka, M. O.; Kusumi, T.; Kakisawa, H.; Kaya, K.; Watanabe, M. M. *J. Am. Chem. Soc.* **1990**, No. 112, 8180–8182.
- (54) Weiz, A. R.; Ishida, K.; Makower, K.; Ziemert, N.; Hertweck, C.; Dittmann, E. *Chem. Biol.* **2011**, *18*, 1413–1421.
- (55) Zhang, Y.; Li, K.; Yang, G.; McBride, J. L.; Bruner, S. D.; Ding, Y. *Nat. Commun.* **2018**, *9*, 1780.
- (56) Roh, H.; Han, Y.; Lee, H.; Kim, S. *ChemBioChem* **2019**, *20*, 1051–1059.
- (57) Okino, T.; Matsuda, H.; Murakami, M.; Yamaguchi, K. *Tetrahedron Lett.* **1995**, *51*, 10679–10686.
- (58) Shin, H. J.; Murakami, M.; Matsuda, H.; Yamaguchi, K. *Tetrahedron* **1996**, *52*, 8159–8168.

- (59) Murakami, M.; Sun, Q.; Ishida, K.; Matsuda, H.; Okino, T.; Yamaguchi, K.
Phytochemistry **1997**, *45*, 1197–1202.
- (60) Fujii, K.; Sivonen, K.; Naganawa, E.; Harada, K.-I. *Tetrahedron Lett.* **2000**, *56*, 725–733.
- (61) Humbert, J.-F.; Barbe, V.; Latifi, A.; Gugger, M.; Calteau, A.; Coursin, T.; Lajus, A.; Castelli, V.; Oztas, S.; Samson, G.; Longin, C.; Medigue, C.; de Marsac, N. T. *PLoS ONE* **2013**, *8*, e70747.
- (62) Lima, S. T.; Alvarenga, D. O.; Etchegaray, A.; Fewer, D. P.; Jokela, J.; Varani, A. M.; Sanz, M.; Dörr, F. A.; Pinto, E.; Sivonen, K.; Fiore, M. F. *ACS Chem. Biol.* **2017**, *12*, 769–778.
- (63) Christiansen, G.; Philmus, B.; Hemscheidt, T.; Kurmayer, R. *J. Bacteriol.* **2011**, *193*, 3822–3831.
- (64) Sano, T.; Kaya, K. *Tetrahedron Lett.* **1995**, *36*, 5933–5936.
- (65) de Carvalho, L. R.; Pipole, F.; Werner, V. R.; Laughinghouse IV, H. D.; de Camargo, A. C. M.; Rangel, M.; Konno, K.; Sant'Anna, C. L. *Braz. J. Microbiol.* **2008**, *39*, 761–769.
- (66) Matsuda, H.; Okino, T.; Murakami, M.; Yamaguchi, K. *Tetrahedron* **1996**, *52*, 14501–14506.
- (67) Welker, M.; Maršálek, B.; Šejnohová, L.; von Döhren, H. *Peptides* **2006**, *27*, 2090–2103.
- (68) Ishida, K.; Welker, M.; Christiansen, G.; Cadel-Six, S.; Bouchier, C.; Dittmann, E.; Hertweck, C.; de Marsac, N. T. *Appl. Environ. Microbiol.* **2009**, *75*, 2017–2026.
- (69) Ishida, K.; Christiansen, G.; Yoshida, W. Y.; Kurmayer, R.; Welker, M.; Valls, N.; Bonjoch, J.; Hertweck, C.; Börner, T.; Hemscheidt, T.; Dittmann, E. *Chem. Biol.* **2007**,

14, 565–576.

- (70) Morita, M.; Hao, Y.; Jokela, J. K.; Sardar, D.; Lin, Z.; Sivonen, K.; Nair, S. K.; Schmidt, E. W. *J. Am. Chem. Soc.* **2018**, *140*, 6044–6048.
- (71) Leikoski, N.; Fewer, D. P.; Jokela, J.; Alakoski, P.; Wahlsten, M.; Sivonen, K. *PLoS ONE* **2012**, *7*, e43002.
- (72) Muro-Pastor, A. M.; Kuritz, T.; Flores, E.; Herrero, A.; Wolk, C. P. *J. Bacteriol.* **1994**, *176*, 1093–1098.
- (73) Cai, Y. P.; Wolk, C. P. *J. Bacteriol.* **1990**, *172*, 3138–3145.
- (74) Seemann, T. *Bioinformatics* **2014**, *30*, 2068–2069.

